https://doi.org/10.31256/Ad3Ko8T

Smart wheelchairs: Visual perception pipeline to improve prediction of user intention

Tomos Fearn
Department of Computer Science
Aberystwyth University, Wales, UK
tof7@aber.ac.uk

Dr. Frédéric Labrosse
Department of Computer Science
Aberystwyth University, Wales, UK
ffl@aber.ac.uk

Dr. Patricia Shaw
Department of Computer Science
Aberystwyth University, Wales, UK
phs@aber.ac.uk

***Abstract*—Wheelchairs aid people with physical disabilities by assisting with mobility, thus improving their independence. Autonomous assistance on wheelchairs are limited to prototypes that provide 'smart functionality', by completing tasks such as docking or terrain adaption. The biggest constraints are navigating within dynamic environments, such as the home.**

This paper describes the data pipeline to automate the wheelchair navigation process, from classifying an object, estimating the user's intention via verbal command (e.g. take me to the fridge) and navigating towards a goal.

Object locations will be registered within a map whilst contextual meta data is calculated. A combination of object classification confidence and object instances is used to calculate the uniqueness of all identifiable objects. Thus, assisting in predicting the user's intention. For example, if a "go to the fridge" request is received, the wheelchair will know that the fridge is located within the kitchen, and therefore drive to the kitchen and then the fridge.

Results show that utilising contextual data reduces the likelihood of false-positive object detections being registered by the navigation pipeline, thus is more likely to interpret the user intention more accurately.

***Index Terms*—robotics, wheelchairs, semantic mapping, ROS**

I. Introduction and Background

Studies show that 40% of wheelchair users with severe physical disabilities require assistance from carers, friends or family to navigate around their environment [1]. This emphasises the importance of introducing a low-cost, robust sensor and software framework that is capable of adapting to the dynamic home environment. We aim to address the challenge of automating the wheelchair navigation process, providing further independence to the user by traversing autonomously towards an intended object.

The platform we use for development is based on a customised differential drive, powered wheelchair chassis. An Arduino Mega is used to interface with the hardware onboard the wheelchair. The main embedded PC uses the Robotic Operating System (ROS) to communicate with hardware and the software stack onboard the wheelchair. The 'RPLIDAR A2' 2D laser scanner, is mounted at the seat height at the front of the wheelchair. The ZED depth camera provides HD RGB images and a pointcloud and is mounted above head height for a person sitting on the wheelchair.

Many recent smart wheelchair prototypes utilise ROS as a middleware, to provide commonly used navigation and localisation tools [2]–[4]. Our novel vision processing pipeline also utilises the ROS ecosystem to communicate between the nodes for calculating user intentions, generating a semantic map, and storing locations of objects on a navigable topological map.

II. Data Pipeline

Due to the absence of sensors onboard the wheelchair to calculate odometry (such as encoders and IMUs), we use RTAB-MAP (Real-Time Appearance Based Mapping) [5] to simultaneously generate a navigable map of the environment and provide a source of odometry. The ZED software development kit is also utilised to calculate visual odometry using feature extraction and loop-closures. A pointcloud of the environment is generated to calculate the coordinates of objects. The LIDAR is a secondary sensor to clean and optimise the local costmap. The 'move_base' ROS package utilises the LIDAR to avoid collisions whilst traversing towards a goal. Its capability to work in low-light conditions and its wider field of view compared to that of the camera makes the LIDAR a safer alternative for avoiding collisions [6].

To detect objects within the environment the Deep Neural Network (DNN) MobileNetV2 [7] object classification is used, trained on the Microsoft COCO dataset [8]. MobileNetV2 requires low computational power (designed for running on a Smart-Phone CPU and interfacing with OpenCV DNN library) and is effective at classifying objects within cluttered images. ROS publishes a custom message per camera frame including the annotated image, *object confidence* and bounding box size. The object confidence (between 0 and 1) is output from the MobileNet (DNN) object recognition node.

To calculate the position of an identified object within a 3D environment, a pointcloud is utilised to calculate the depth of the centre pixels of a bounding box. RTAB-MAP generates a navigable map of 3D coordinates for the wheelchair. Translating the point coordinates from the camera to the global map layer occurs using the ROS transform library. Objects are uniquely identified with an ID and name, a transform is placed on the corresponding coordinates on the map.

A. Training

The smart wheelchair must be trained once by the operator when introduced to a new home environment. When entering a new room, a label must be provided (i.e. kitchen, bedroom, lounge) (Fig. 2). A ROS transform is generated using the pose

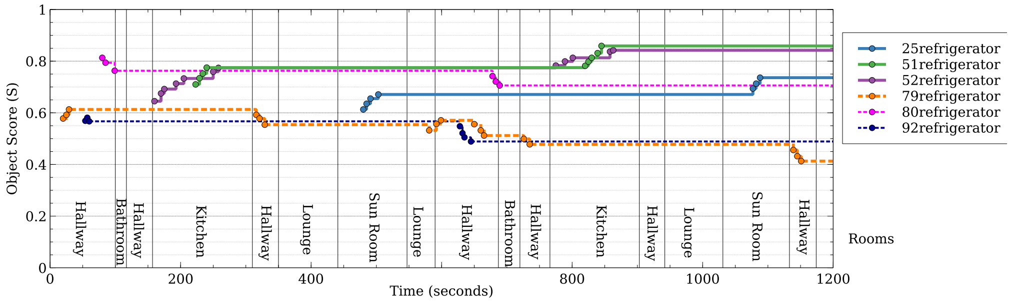

Fig. 1. Object Context Score over time

of the wheelchair and is placed on the global map. All new objects identified within the room are assigned a room label.

Fig. 2. Top-down view of the home environment including correctly classified refrigerator objects circled green, with false-positives circled red

B. Calculating context of an object

One of the primary issues within this methodology, is the presence of false-positive detections classified by the DNN object detection (e.g. classifying a white wall incorrectly as a refrigerator). A novel method was devised to overcome this issue (the 'Wheelchair_Context' ROS package [9]), by analysing the context (*object score*), *object confidence* (from the MobileNet (DNN) object recognition node), *object weighting* (likelihood of object being detected in previous location), *object uniqueness* (the total number of objects found in the home environment). To prevent the object score becoming too big or small (thus affecting the likelihood of navigating towards an object), object score is bounded between 0 and 1.

Fig. 1 visualises the change in *object score* over *time (secs)* as the wheelchair traverses between rooms. Three correctly detected refrigerators (object ID 25, 51 and 52) and three false-positive object detections (object ID 79, 80 and 82) are plotted.

On revisiting locations, re-detection increases confidence, with the real objects being reliably re-detected. The result is false-positives becoming less influential and navigation takes the user to the intended object.

C. Estimating the user's intention

Our navigation package consists of three possible navigation modes, based on how specific the user's instructions are:

1) No room named, but object name is supplied — based on all objects within the environment, which object is most likely the user's intention?
2) Both room and object names are supplied — list of potential objects to navigate to is significantly shortened due to associated rooms.
3) Only the room name is provided — the wheelchair will navigate to the specified room transform, whilst avoiding any objects that may be obstructing the goal.

III. Conclusion

We have presented an overview of an autonomous wheelchair capable of creating a semantic map of objects and rooms, whilst calculating the user's intention. The method provides a low-cost solution to reducing the influence of false-positive DNN object detections in predicting the user's intention. Our 'object-room tag' vision-only approach provides the flexibility of operating in either an open plan or distinct room house. Future work also includes providing customised approach behaviours for individual objects within the home (e.g. slowly approaching a table and docking the wheelchair).

References

[1] L. Fehr, W. Langbein, and S. Skaar, "Adequacy of power wheelchair control interfaces for persons with severe disabilities.," *J. of Rehabilitation Research and Development*, vol. 37, no. 3, pp. 353–360, 2000.

[2] Z. Li, Y. Xiong, and L. Zhou, "ROS-Based Indoor Autonomous Exploration and Navigation Wheelchair," *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 2, 2017.

[3] L. Lecrosnier, R. Khemmar, N. Ragot, B. Decoux, R. Rossi, N. Kefi, and J. Ertaud, "Deep Learning-Based Object Detection, Localisation and Tracking for Smart Wheelchair Healthcare Mobility.," *International Journal of Environmental Research and Public Health*, vol. 18, 2021.

[4] H. S. Grewal, N. T. Jayaprakash, A. Matthews, C. Shrivastav, and K. George, "Autonomous wheelchair navigation in unmapped indoor environments," in *I2MTC*, pp. 1–6, 2018.

[5] M. Labbé and F. Michaud, "RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," *Journal of Field Robotics*, vol. 36, 2019.

[6] T. Fearn, F. Labrosse, and P. Shaw, "Wheelchair navigation: Automatically adapting to evolving environments," in *TAROS*, pp. 496–500, Springer, 2019.

[7] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," 2017.

[8] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, "Microsoft COCO: Common Objects in Context," 2015.

[9] T. Fearn, "Wheelchair Context ROS Package." https://github.com/fearn-robotics/wheelchair_context, 2022.